Genetic Maps/1984, volume 3: *A Compilation of Linkage and Restriction Maps of Genetically Studied Organisms*. Edited by STEPHEN J. O'BRIEN. New York: Cold Spring Harbor Laboratory. 1984. 583 pages (approx.). $28 U.S.A. $33.60 elsewhere. ISBN 0 87969 171 9.

One can only review a compilation like this by presenting a mini-compilation, to give the prospective reader an idea of what to expect. I have to say first, however, that leafing through this handsome book puts me in mind of the Spanish Nobles who were viewing the great Mosque in Seville after driving out the Moors, in 1401. Having determined to build a Christian Cathedral on the ruins of the Mosque, they said 'Let us build such a great edifice as all men who see it shall think us mad'. Stephen O'Brien has encouraged (or found) the same adventurous spirit in the numerous authors who have contributed to this large volume and, like the builders of Seville Cathedral, they deserve much congratulation. One might add that there is also a suggestion of the many tongues on the Tower of Babel in both the variety of languages used to symbolize genes in different species and the variety of typewriter and computer print-outs used for gene lists and texts. None of these is a serious cause for concern, as one can get used to any gene language if one has to, even when such a variety of symbols as 7, 80, 962, *f*, rd^2 and rib^2 occur in the gene list for a single organism. One printing defect in some sections which occurs also in many manuscripts, should be noted, however. This is the use of the same typewriter symbol for numeral '1' and lower case letter 'el' – as an editor, I cannot always guess correctly which is intended, and even less can our printer.

Genetic Maps was the result of a bold plan, at the 1979 Human Gene Mapping Conference in Edinburgh, to extend the proposed listing of human genes to include the maps of all species that had been studied, the total lists to be published in one volume. The collection was to be revised every 2 years, making the previous volumes obsolete. So volumes 1 and 2, which appeared in 1980 and 1982, can now be discarded in favour of volume 3, containing data mainly up to the beginning of 1984.

An indication of recent progress can be obtained by comparing Volumes 2 and 3. Volume 2 contained assorted maps for man and eight other primates, seven other vertebrates, two insects, one nematode, five fungi, two protozoa, one alga, four plants, fourteen bacteria, six bacteriophages and six animal viruses – a very miscellaneous sample of living forms which reflects the peculiar interests of geneticists and the limitations of their material. The new volume has added the owl monkey, gibbon, cow, pig, horse, Chinese hamster and Chinese hamster ovary cell line, *Rana pipiens*, two groups of fish (Poeciliid and Salmonid), one bacterium (*Proteus morganii*) and twenty-five animal retroviruses; and it also gives details of the Protein and Nucleic Acid Sequence Databases, of which more later.

A few of the maps are virtually or entirely unchanged in the new volume, in particular *Bacillus subtilis*, *Pseudomonas putida*, *Neisseria gonorrhoeae*, the three *Rhizobium* species *meliloti*, *phaseoli* and *trifolii* (because the genes of current interest in these Rhizobia are mainly plasmid borne), *Acinetobacter calcoaceticus* and *Podospora anserina*; but generally much effort has gone into bringing the information up-to-date, often by scavenging for unpublished results. Gene and restriction site maps of mitochondrial DNA are given for *Saccharomyces cerevisiae*, *Aspergillus nidulans*, *Neurospora crassa*, *Homo sapiens* and *Chlamydomonas rheinhardtii* (for which chloroplast DNA is also included). Restriction site maps are also given for the six phages (Lambda, ϕX174, T4, P1, P2 and P4), Polyoma Virus, SV40, Hepatitis B virus, a number of Human Adenovirus serotypes, and the twenty-five retroviruses, but not for herpes simplex virus DNA, whose 150000 base pairs are doubtless over-long for restriction analysis. By far the largest section is the 46 pages on the human genome, which includes eight gene maps and lists with much explanatory information attached. This forms a key reference work in itself, and is bound to be of great interest to a wide variety of readers.

Many of the gene maps are largely or entirely confined to biochemical loci, e.g. Chinese hamster (40 loci), cat (40 loci of which only one is a colour gene), dog (35 loci), pig (38 loci) and cow (32 loci). Some of these species must possess known colour or morphological mutants, as is the case with the horse (17 loci include two for coat colour), rabbit (53 loci include 12 affecting coat colour or fur characters), and the chicken (92 loci, either on a linkage group or linked to at least one other gene, include a number which are not biochemical by any normal definition). The primate maps (excluding man) are all based on biochemical markers and range from 42 loci in the chimpanzee to 19 in the African green monkey.

It is also of some interest to compare the numbers of loci listed for different species. Man is well ahead of the field, with 3577 loci (of which about half were not yet fully identified or confirmed at 1st February 1984). *Drosophila melanogaster* may possibly exceed this number, but we shall not know until the revised Lindsley and Grell listing (by Lindsley, Grell and Zimm), which is expected shortly, is available. In the present volume we find, for *D. melanogaster*, lists of biochemical loci, cloned DNA segments, and transfer RNA and U-RNA in situ hybridisation data. The mouse (1068 loci including 364 not yet mapped) has overtaken *Escherichia coli* (1026 loci), which leaves its rival *Salmonella typhimurium* (nearly 600 loci) well behind. But to have identified some 30–40 % of the genes of *E. coli* is undoubtedly a triumph, to which the *umuC* gene (not, apparently, present in all bacteria) has doubtless made a valuable contribution, Several species have around 500 loci identified: *Neurospora crassa* 525, *Aspergillus nidulans* 485, *Saccharomyces cerevisiae* 412, maize 479, *Caenorhabditis elegans* 446, and tomato 328 (if my gene counting has been accurate). I was astonished to find that there are 24 different loci in the tomato giving resistance to *Cladosporium fulvum*.

New to Volume 3 are details of the Protein and Nucleic Acid Sequence Databases, and how to obtain any particular sequence of interest and all relevant information about it. The NBRF-PIR Protein Sequence Database contains over 500000 residues in over 2600 entries, which include amino-terminal and substantial fragmentary sequences, as well as all completely sequenced proteins, amino acid sequences translated from nucleic acid sequences, together with hypothetical proteins from open reading frames. A condensed list of these sequences, divided into 813 items or subgroups, gives one a very good indication of the weight of knowledge that has been collected in this field. It is obviously much easier to obtain the DNA sequence and convert this into amino acids, and it would be interesting to know how many of these 2600 sequences were actually obtained directly from proteins. Those not up to date in proteins sequenced may like to know that the complete haemoglobin beta chain sequences include 61 mammals, 8 birds, 4 amphibians, 3 reptiles and 3 fishes. One can retrieve sequences from this Database in a number of ways, too numerous to list here but set out on page 522, which show the care and ingenuity which have gone into designing the retrieval programmes.

The Nucleic Acid Sequence Databases include those of NBRF, listed here by organism and gene of that organism, with a separate listing of functional RNAs, and the Nucleotide Sequence Data Banks of GenBank and EMBL, which cooperate closely and differ only in details of format. The 4286 sequences in the GenBank Database at 14/2/1984 are listed here under their 9-or-fewer character names, together with an explanation of what each sequence consists of. The character name is used to obtain the sequence through a suitable computer link. This database contained over 2·8 million bases at February 1984, and will probably exceed 4 million when this review is published. Looking through these lists, one notes that about 212000 base pairs of *E. coli* have been sequenced, or about 5% of the total *E. coli* DNA. Sequencing has proceeded nothing like so far as this in any other prokaryote or eukaryote, so clearly many decades (centuries?) of rewarding labour lie ahead for molecular geneticists.

This book is altogether a very impressive, valuable and even readable work of reference, which is remarkably cheap for the amount of up-to-date information it

contains. It should obviously be on the shelves of every biological library, since no biologist can escape from genetics today. Dipping into it makes one aware how thinly genetic knowledge is spread over the living world, and leaves one convinced that the total genes yet named or identified (surely less than 10000), and the 3–4 million nucleotides sequenced add up to only a bucketful in the lake – and this is good news for those who can still lay their hands on money for research. The book is also of value in enabling one to find out quickly what the boys and girls in the other back rooms have been doing.

I hope we can look forward to the next volume in 1986, though it will need even more of the Seville spirit than this one. If it goes into production, may I offer a few suggestions to the editors and contributors; (i) it would be very helpful for page shufflers to have the name of the organism given on each page. (ii) there are a number of organisms with developing (or complete) gene maps which might find a place: phage Mu (which I sadly missed), phage T7 (with its complete genome already sequenced), perhaps P22, a few more Enterobacteria such as *Klebsiella* (an early map badly needs up-dating before it is published) and possibly *Erwinia* (interest in growing on certain species), *Physarum polycephalum* to set against *Dictyostelium discoideum* (the Physarians claim *Dictyostelium* is not a true slime mould), and even other species of Drosophila. (iii) I would have found a little more text with each organism helpful, particularly if it gave some information on related organisms not included and where one could look for genetic data on them. This would help those who wanted a wider view of, say, bacteriophage genes. For example, the complete nucleic acid sequences of f1, fd, g4, m13 and ms2, and t7, as well as the listed lambda and phiX174, are available. As a footnote, I met a few puzzles when looking through the Genbank Database list. Where is influenza virus? Ah, of course, it is coded under 'fl' for flu! Under bacteria, I found sau (clearly *staphylococcus aureus*), sdy (*Shygella dysenteriae*) and sma (*Serratia marcescens*), but I was stumped by *S. pneumoniae* and *S. fradiae*. Perhaps if you ask the computer it will tell you what these organisms are. Or perhaps everyone else knows.

ERIC REEVE
Institute of Animal Genetics
University of Edinburgh

Modern Approaches to Vaccines. Edited by ROBERT CHANOCK and RICHARD LERNER. Cold Spring Harbor Laboratory. 1984. 500 pages (approx.). $52.00 U.S.A. $62.40 elsewhere. ISBN 0 87969 165 4

This book is the product of a meeting held in Cold Spring Harbor in 1983, which was to be the first in a series of meetings discussing modern approaches to vaccines. This first meeting aimed to cover the topics of the molecular and chemical basis of virus virulence and immunogenicity. The sixty-six papers presented at the symposium are reproduced in the book, thus bringing together in a single volume a timely commentary on the evolving concepts of viral immunogenicity and on the applications of recent advances in biotechnology to the production of viral vaccines. A rapid leaf through the pages encourages the reader to a sense of great optimism for a future in which safe and effective control of viral infection will be based at last on scientifically proven principles of virus virulence and immunogenicity.

The headings of the five sections into which the presentations are divided indicate the attempted plan of the conference proceedings, but there is overlap and repetition between sections and three papers on bacterial vaccines have been included. As might be expected the largest number of papers are presented in Section 3 which covers the topics of Cloning and Expression of Viral Genes. Papers on picornaviruses (poliovirus and foot and mouth disease virus, FMDV) Hepatitis B virus and influenza virus appear regularly throughout the book, but other important virus infections for which effective vaccines are also desperately needed are given less prominence although there are four papers each on rabies virus and herpes simplex virus.